10/12/2022

Data Warehouse

Team [3]

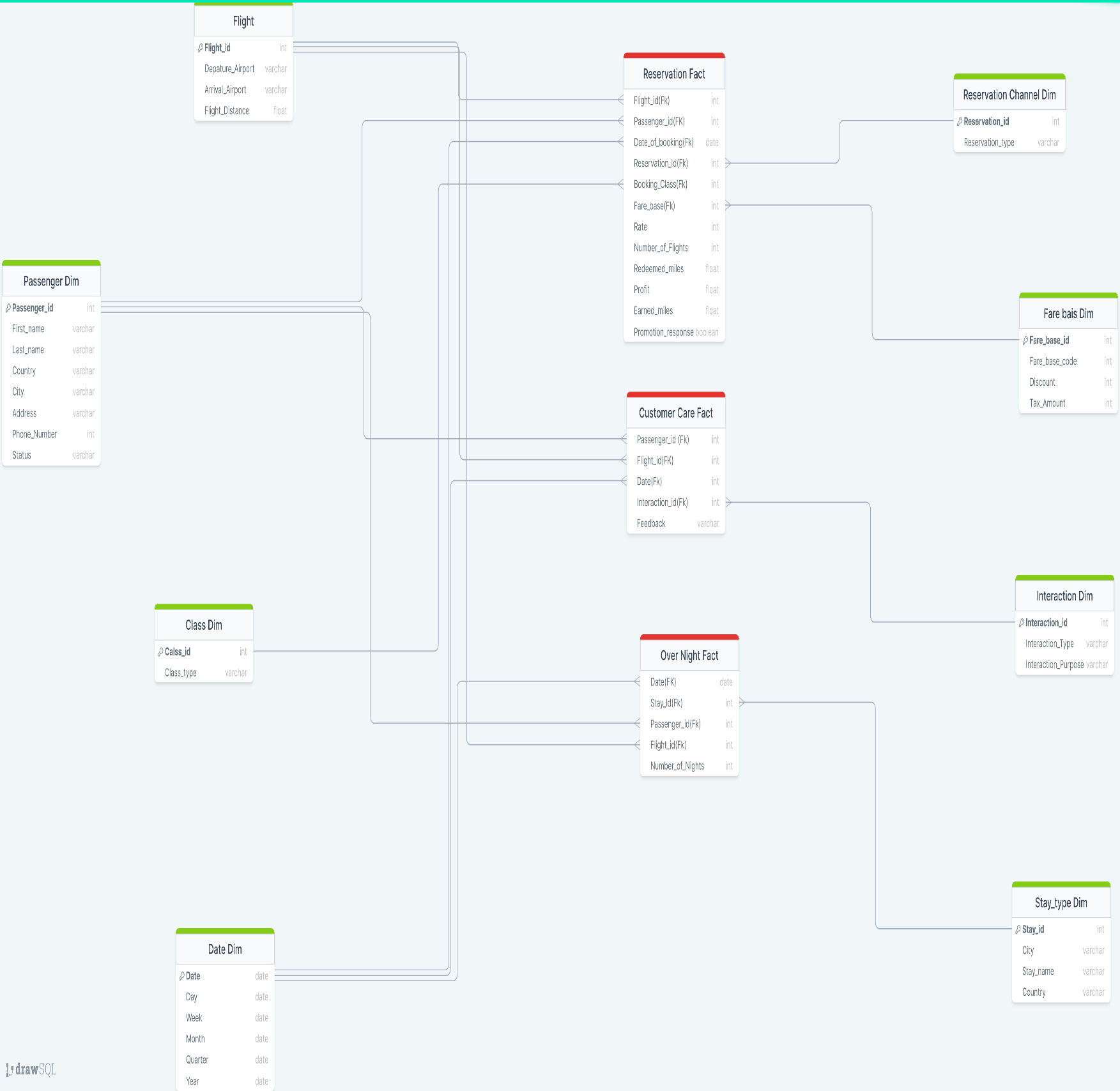


1. **Bus Matrix:**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Facts\Dimensions | Flight | Passenger | Class | Date | Reservation channel | Fair base | Interaction | Stay type |
| Overnight | **✓** | **✓** |  | **✓** |  |  |  | **✓** |
| Reservation | **✓** | **✓** | **✓** | **✓** | **✓** | **✓** |  |  |
| Customer care | **✓** | **✓** |  | **✓** |  |  | **✓** |  |

1. **Logical Model**

**Note: This model has been built using drawsql (Web application)**

**** [**https://drawsql.app/teams/dwh-1/diagrams/dwh-2**](https://drawsql.app/teams/dwh-1/diagrams/dwh-2)

1. **More details about Logical Model**

* Business Case:

Analyze the current business processes and expand the company by discovering new opportunities

* Granularity:

Atomic per reservation

Daily

* Dimensions:

Flight\_dim (#flight\_id, departure\_airport, arrival\_airport, flight\_distance)

Flight\_dimension: is essential to describe all used fact tables

flight\_id: acts as a primary key for the table

departure\_airport: to describe the primary key attributes

arrival\_airport: to describe the primary key attributes

flight\_distance: is an important attribute in the case study

Passenger\_dim (#passenger\_id, first\_name, last\_name, country, city, address, phone\_number, status)

Passenger\_dim: to describe the main profitable unit in the business

Passenger\_id: acts as a primary key

First\_name:

Last\_name:

Country: essential to make location analysis

City: essential to make location analysis

Address:

Phone\_number: to be able to reach customer

Status: customer segmentation is one of the most important attributes regarding business analysis to help business and marketing campaigns

Class\_dim (#class\_id, class\_type)

Class\_id: acts as a primary key

Class\_type: to describe which package the customer purchases

Date\_dim (#Date\_key, Date\_A, year, month, week, day)

Date\_dim: is used in all tables to be able to make timeline analysis and historical analysis

Date\_key: acts as a primary key

Date\_A: the real dates in the dataset

Year:

Month:

Week:

Day: satisfying the level of granularity

Reservation\_channel\_dim (#reservation\_id, reservation\_type)

Reservation\_channel\_dim: essential for investing and making offers

Reservation\_id: acts as a primary key

Reservation\_type: essential for financial team analysis

Fare\_base\_dim (#fare\_base\_id, fare\_base\_code, discount, tax\_amount)

Fare\_base\_dim: its an international reservation classification

Fare\_base\_id: acts as a primary key

Fare\_base\_code: to show the type of reservation features and details

Discount: related to the fare base code

Tax\_amount: also is known as per fare base code

Interaction\_dim (#Interaction\_id, interaction\_type, interaction\_purpose)

Interaction\_dim: to describe the type of customer care service and improve the company services

Interaction\_id: acts as a primary key

Interaction\_type: (face to face, call center support, etc..)

Interaction\_purpose: (complain, inquiry, etc….)

Stay\_type\_dim (#stay\_id, city, stay\_name, country)

Stay\_type\_dim: to make analysis on the company customers to explain a probable area of investment and business extension

City: location for analysis

Stay\_name: the type of stay is very important for targeting a new business

Country: location for analysis

* Facts:

We found many different analysis areas; therefore, a lot of fact tables have been generated using **(Galaxy Schema)**

Reservation\_fact:

**Profit**: one of the most important KPIs for the business success and very essential to evaluate different dimensions of business

**Number of flights**: to check the passengers flights quantity to widen the area of analysis

**Rate**: based on the case study the business gives a high priority to any type of feedback coming from his customer to improve his service

**Earned miles:** an essential measure to track the customer usage to improve customer loyalty

**Redeemed miles:** an essential measure to track the customer usage

**Promotion response**: the business needs to measure the success of the marketing campaigns

Customer\_care\_fact:

**Feedback**: the case gives a very high priority to the customer’s feedback to enhance the business so creation of a measure for the feedback is not a choice

Over\_night\_fact:

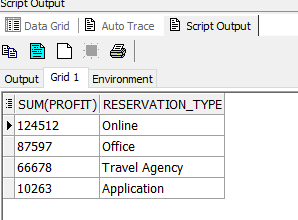
**Number of overnights**: In order to widen the area of investment, the overnights is very essential measure analyze and generate new opportunities

1. **Logical data model to physical data model (Done)**
2. **& 6- SQL Queries**

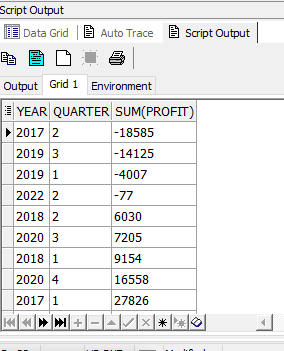
* select sum(profit), reservation\_type

from reservation\_channel\_dim , reservation\_fact

where reservation\_channel\_dim.reservation\_id = reservation\_fact.reservation\_id

group by reservation\_type;

* select year, quarter, sum(profit)

from reservation\_fact, date\_dim

where reservation\_fact.date\_of\_booking = date\_dim.date\_key

group by year, quarter

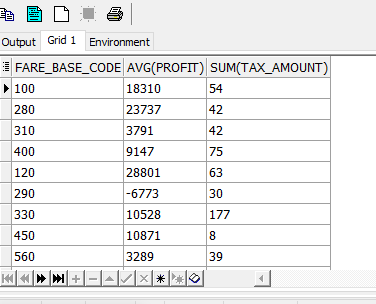
order by sum (profit);

* select fare\_base\_code, avg(profit), sum(tax\_amount)

from reservation\_fact, fare\_bais\_dim

where reservation\_fact.fare\_base = fare\_bais\_dim.fare\_base\_id

group by fare\_base\_code;

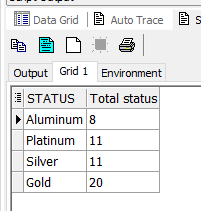


* select status, count(status) as "Total status"

from passenger\_dim

group by status

order by count(status);



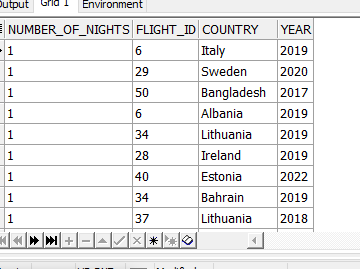
* SELECT onf.number\_of\_nights, onf.flight\_id, std.country, dd.year

FROM stay\_type\_dim std, over\_night\_fact onf, date\_dim dd

WHERE std.stay\_id = onf.stay\_id

AND onf.date\_key = dd.date\_key

ORDER BY onf.number\_of\_nights;



Thank You

Ali Mostafa

Hassan Essam

Khaled Magdy

Mazen Magdy

Mohannad Riffat